On optimal tax differences between heterogenous groups

Klaus Beckmann
Helmut Schmidt University
Department of Economics
Holstenhofweg 85
22043 Hamburg
Germany
klaus.beckmann@hsu-hh.de

Andrea Schneider*
University of Münster
Department of Economics
Wilmergasse 6-8
48143 Münster
Germany
andrea.schneider@uni-muenster.de

November 25, 2012

Abstract

This paper considers optimal linear tax structures that are differentiated according to group membership. Groups can be heterogenous with respect to both preferences and abilities. Contrary to most arguments in favour of tax privileges for certain groups, e.g. gender-based taxation, it is shown that consideration of the first moment of the relevant distributions (the average labour supply elasticity of the groups) is insufficient. We discuss the factors on which efficient differentiation would depend and argue that the results of our analysis militate for adopting horizontal equity as a rule of thumb.

JEL classification : H21; D31

Keywords : optimal linear income taxation; preference heterogeneity; horizontal equity; gender-based taxation; tax privilege

*Corresponding author. Phone +49-251-83-22823. Fax +49-251-83-22826.

1 Horizontal equity and efficient taxation

Horizontal equity, viz. the social norm that equal individuals be treated alike, requires that the treatment of persons depend only on their individual properties.[1] Social privilege, on the other hand, allows for the treatment of individuals to be differentiated according to group membership, such that *properties of the group* can also be taken into account. It is precisely in this sense that the French Revolutionaries wanted privileges to be abolished, demanding *égalité* – i.e., equality before the law.[2]

It is quite easy to show that horizontal equity is incompatible with the Pareto principle. Sen's (1970) seminal paper was probably the first to develop an argument in this direction, a general version of which is given in lemma 1 below.[3]

LEMMA **1. Conflict between absolute goals and the Pareto principle.** *Let social welfare W depend on the life satisfaction u_i of all n individuals in society (welfarism) and some other objective g.*

$$W = W(u_1, \ldots, u_n; g) \tag{1}$$

If individual satisfaction is independent of g for all individuals $i \in \{1 \ldots n\}$, then the Pareto Principle implies that society must not accord any weight at all to g.

Of course, this still leaves room for a Paretian to justify horizontal equity as an *instrumental* goal by allowing for $\frac{\partial u_i}{\partial g} \neq 0$ to hold for at least one individual i. Constitutional lawyers, on the other hand, would probably join the horizontal equity camp on principled grounds and eschew the Pareto principle altogether.

[1]On the concept of horizontal equity, see Musgrave (1990).

[2]Notions of communality, including nascent welfare statism, were captured by the third watchword, "fraternité". See Henking (1989) for a careful treatment of these issues as well as for a comparison to the U.S. constitutional tradition.

[3]A proof of lemma 1 is as follows: Totally differentiate (1) to find

$$\mathrm{d}W = \sum_{i=1}^{n} \frac{\partial W}{\partial u_i} \mathrm{d}u_i + \frac{\partial W}{\partial g} \mathrm{d}g$$

Note $\frac{\partial u_i}{\partial g} = 0$ by assumption as the absolute goal g does not affect people's utilities directly. Recall that the (weak) Pareto principle can be stated as follows: $\mathrm{d}u_i \geq 0 \, \forall \, i \in \{1 \ldots n\} \Rightarrow \mathrm{d}W \geq 0$. It is now obvious that this implies $\frac{\partial W}{\partial g} = 0$.

Economists are more of a Paretian persuasion, and have used some version of lemma 1 to demonstrate that a neglect of horizontal equity can lead to a Pareto improvement. A seminal example is Stiglitz' (1982) paper on optimal random taxation. Recent examples include an argument by Alesina, Ichino and Karababournis (2007) that females ought to be taxed more lightly through the income tax than males because their *average* elasticity of labour supply is demonstrably lower.

We contend that arguments of this ilk are seriously incomplete. As the excess burden of taxation is non-linear in the tax rate, a comparison of first moments is clearly insufficient to justify differentiating tax rates (treatments more generally) among groups. We provide an example where the group with the smaller average elasticity of labour supply should be taxed more heavily, and also provide some more general results demonstrating the factors on which an efficient differentiation of tax rates would depend.

This also opens up another non-instrumental angle for justifying horizontal equity (the rejection of privileges for certain groups, such as females): introduction of a Pareto-optimal tax policy can involve a partitioning of tax payers into sub-groups too small to be justified given the administrative costs of doing so, and given the inevitable vagaries of defining multiple sub-groups (Congleton 1997).

Differentiated treatment is not a new topic in the literature; for example, Brander and Spencer (1985) already provided a formal model to discuss the impact on the Ramsey rule of introducing heterogeneous groups of customers. Sandmo (1993) also discussed optimal taxation with heterogenous preferences.

More recently, Lockwood and Weinzierl (2012) analysed the impact of heterogeneous preferences on optimal taxation, finding that this would under fairly general conditions call for *less* redistribution than in the standard homogenous preference case. Their paper starts from a standard Mirrlees model and add additional dimensions of heterogeneity. This differs from the approach taken here in that we focus on the taxation of *groups*, which may differ in the distribution of parameters in the various dimensions. We also explicitly consider the properties of these distributions.

We also diverge from the existing literature in that we examine *differentiated linear tax structures*, which consist of a separate marginal tax rate for each group and a lump-sum tax component that is uniform across groups. In a sense that will be made clear below, this can be interpreted as a part of a larger optimal tax problem, in which the different treatment of (small) groups of the same ability is analysed, all else being held equal.

Section 2 sets out the model, while sections 3 and 4 prove the main results. Section 5 concludes.

2 The model

Assume a fixed mass of individuals who derive benefit from consumption and leisure time. Moreover that taxpayers can be partitioned into n groups of unit mass, so that an individual taxpayer i of group j is indexed by ij where $i \in [0,1]$ and $j = \{1, \ldots n\}$. Taxpayer groups are in all respects identical, except being characterized by different and independent distributions of the preference for leisure time, α_{ij}, and (possibly) different wage rates. The distribution of the preference parameter for taxpayer group j is given by $f_j(\alpha_{ij})$.

In this paper we will analyse optimal tax rates when groups can be taxed differently while their is no differentiation within groups. The government can set a group-specific linear income tax schedule with the marginal rate t_j and a lump-sum component T.

Each individual i of group j maximises its quasi-linear utility function

$$u_{ij}(c_{ij}, l_{ij}) = c_{ij} + \alpha_{ij}\sqrt{l_{ij}} \tag{2}$$

over the choice of consumption, c_{ij} and leisure time, l_{ij}, where the entire time budget is normalized to unity. The wage rate (labour productivity) w_j is exogenously given and can vary across groups, but not within groups.

As there are no income effects in our model, we can ignore T when computing the optimal labour-leisure choice. When labour income is taxed proportionally at the rate t_j, individual's budget constraint is

$$c_{ij} \leq (1 - t_j)(1 - l_{ij})w_j \tag{3}$$

This simple problem affords a closed form solution for leisure and consumption

$$l^*_{ij} = \frac{\alpha^2_{ij}}{4(1-t_j)^2 w^2_j}\,, \quad c^*_{ij} = (1-t_j)w_j - \frac{\alpha^2_{ij}}{4(1-t_j)w_j} \tag{4}$$

where individual i enjoys indirect utility

$$v_{ij}(w_j, t_j) = (1 - t_j)w_j + \frac{\alpha_{ij}^2}{4(1 - t_j)w_j}. \tag{5}$$

Note that this increases in the weight α of the non-linear preference for leisure, so individuals who are relatively less interested in consumption are better off in this framework, *ceteris paribus*. From the solution (4), we can easily compute the (compensated) elasticity of labour supply with respect to the net wage as

$$\eta_{ij} = \frac{\partial(1 - l_{ij}^*)}{\partial w_j(1 - t_j)} \cdot \frac{w_j(1 - t_j)}{1 - l_{ij}^*} = -\frac{2\alpha_{ij}^2}{\alpha_{ij}^2 - 4(t_j - 1)^2 w_j^2}. \tag{6}$$

Taking the first derivative of (6) with respect to α_{ij}, we see immediately that η_{ij} is strictly increasing in α_{ij}. In the remainder of this paper, we will therefore argue in terms of α_{ij} instead of the elasticity of labour supply itself.

To calculate the excess burden caused by income taxation of individual ij, we need to compare the revenue from the proportional tax on labour income

$$R_{ij} = t_j w_j \big(1 - \frac{\alpha_{ij}^2}{4(1 - t_j)^2 w_j^2}\big). \tag{7}$$

to the revenue generated by a lump-sum tax keeping the individual at the identical level of indirect utility v_{ij}. With a lump-sum tax Θ_{ij}, the individual's problem becomes

$$u_{ij}(c_{ij}, l_{ij}) \rightarrow \max! \qquad \text{s.t.} \quad c_{ij} \leq (1 - l_{ij})w_j - \Theta_{ij} \tag{8}$$

with the closed form solution

$$l_{ij}^{**} = \frac{\alpha_{ij}^2}{4w_j^2}, \quad c_{ij}^{**}(T_{ij}) = w_j - \frac{\alpha_{ij}^2}{4w_j} - \Theta_{ij}. \tag{9}$$

Using (9) and (5), we find the lump-sum tax yielding the utility v_{ij} to be

$$\Theta_{ij} = t_j w_j - \frac{\alpha_{ij}^2 t_j}{4(1 - t_j)w_j} \tag{10}$$

and can finally compute the deadweight loss D_{ij} arising through the taxation of ij's labour income:

$$D_{ij} = \Theta_{ij} - R_{ij} = \frac{\alpha_{ij}^2 t_j^2}{4(1-t_j)^2 w_j}. \tag{11}$$

Note that D_{ij} increases in the parameter α_{ij} and, therefore, the labour supply elasticity η_{ij}.

The overall deadweight loss, D_j and tax revenue, R_j, generated from the jth group (out of n) can then be found by simply integrating (11) and (7), respectively. This gives

$$R_j = \int \Big(t_j w_j - \frac{\alpha_{ij}^2 t_j}{4(1-t_j)^2 w_j} \Big) f_j(\alpha_{ij}) \mathrm{d}\alpha_{ij} \tag{12}$$

and

$$D_j = \int \frac{\alpha_{ij}^2 t_j^2}{4(1-t_j)^2 w_j} f_j(\alpha_{ij}) \mathrm{d}\alpha_{ij}. \tag{13}$$

We impose a balanced government budget and analyse a *differentiated linear tax structure* – which consists of a separate marginal tax rate per group and a uniform lump-sum compenent T – that minimises the overall deadweight loss subject to the budget constraint. We obtain the following well-behaved non-linear minimisation problem:

$$\min_{t_1,\dots t_n} \sum_{j=1}^{n} D_j(t_j) \quad s.t. \sum_{j=1}^{n} R_j(t_j) \geq T. \tag{14}$$

In solving (14), we treat T (i.e., the lump-sum subsidy for everyone) as exogenous and minimise *unweighted* deadweight loss. This means that we effectively assign equal money metric distributional weights – $\frac{\partial W}{\partial u}\frac{\partial u}{\partial c}$'s in terms of equation (1) – to all individuals and, because they have equal mass, to all groups. At first blush, this appears to be a rather restrictive assumption. It can, however, be justified in one of the two following ways:

1. As individual preferences (2) imply transferable utility in our model, the problem (14) is in fact compatible with maximising a utilitarian

social welfare function.[4] Its solution determines the optimal structure of per-group t_j^*'s, given a target subsidy T^*. While we cannot endogenously determine the latter using the solution to (14), this is rather immaterial for our point about the optimal structural differentiation between groups.

2. Consider a problem where the overall redistributiveness of the tax system has been determined beforehand, and the only issue in question is the efficient treatment of two groups that are similar in terms of their ability. For concreteness, imagine that we are discussing possible tax relief for female professionals (as opposed to males in the same occupation). If the two groups in question are relatively small and about equally well-off, the sub-problem can then be couched in terms of equation (2).

In general, it is not possible to give a closed form solution for this minimisation problem without introducing additional assumptions concerning the distribution $f(\cdot)$.

3 First moments do not suffice: a counter-example

We begin by demonstrating that optimal labour income tax rates do not only depend on the *average* labour supply elasticity of groups, which is normally focused on in the application of optimal taxation theory to policy, but also on the heterogeneity within each group. Proposition 1 summarises this point.

PROPOSITION **1.** *A higher mean elasticity of a group is neither necessary nor sufficient for the optimal marginal tax rate for this group to be lower than for the other group.*

Proof. Our proof of proposition 1 is by counter-example for two groups $j = \{1, 2\}$ and uniformly distributed preferences, i.e. $\alpha_{ij} \sim \mathcal{R}[\mu_j - \sqrt{3}\sigma_j, \mu_j + \sqrt{3}\sigma_j]$. Fixing wage rates at $w_1 = w_2 = 2$ and $T = 1$, table 1 shows the

[4] Formally, the marginal utility of income is the same for all individuals and equal to $\frac{\partial u_{ij}}{\partial c_{ij}} = 1$ given the assumptions of our model. For a classical utilitarian, $\frac{\partial W}{\partial u_{ij}} = 1$ for all ij, and distributional weights only depend on the marginal utility of income.

optimal tax rates for various combinations of the means μ_1, μ_2 and standard deviations σ_1, and σ_2:

	μ_1	σ_1	μ_2	σ_2	t_1^*	t_2^*
Baseline	0.5	0.1	0.5	0.1	0.258	0.258
Increase in μ	**0.6**	0.1	0.5	0.1	0.236	0.282
Increase in σ	0.5	**0.2**	0.5	0.1	0.251	0.265
"Counter-example"	0.5	0.1	**0.45**	**0.35**	0.273	0.244

Table 1: Simulation results

The first line in table 1 establishes a baseline case, while the second shows the well-known effect of an increase in the *average* elasticity of labour supply for a group: the optimal group-specific tax rate goes down, while the other group is taxed harder. The third line confirms our intuition that an increase in the variance of the elasticity for a group, other things being equal, will lead to lower taxation of this particular group. The reason for this is the non-linearity of the deadweight loss; adding to the tail ends of the distribution will entail adding cases with a disproportionally higher excess burden on the right, which outweighs the reduction through the additional cases on the left of the median.

Finally, the fourth line demonstrates, though only by way of example, that the second effect discussed above can dominate the first. Even though group 2 exhibits a smaller elasticity of labour supply on average, its tax burden is lower, on account of the greater standard deviation in the distribution of the parameter α. It is possible, then, for a group with a *lower average elasticity of labour supply to be taxed less heavily in an optimum.* Clearly, a look at the first moment of the distribution alone does not suffice. □

Note, that in the example the labour productivity is identically distributed in each group. Thus, the Mirrleesian optimal taxation theory would not call for any different treatment of the groups at all. In addition, the optimal differentiation of tax rates between groups which we derive clearly violates the "preference neutrality" requirement imposed in Lockwood and Weinzierl (2012), which states that in the absence of ability differences no redistribution should take place at all. We shall return to this point in the discussion.

4 Optimal differentiation of tax rates across groups

Consider the impact of a change in a moment of $f(\cdot)$ on the optimal tax rates, all other things being equal. Intuitively, it seems plausible to conjecture (cf. also table 1) that both an increase of the first and second moment for a group would lead to a *lower* tax rate for that group in the optimum. Given T and the independence of group distributions, we would also expect the tax rates for all other groups to rise, albeit to a different degree. This conjecture, however, would be premature, as proposition 2 shows.

PROPOSITION **2.** *Minimising the total (unweighted) excess burden of taxation through differentiating marginal tax rates across heterogenous groups, the optimal marginal tax rate for group j, $j \in \{1, \dots, n\}$, is*

(i) decreasing in the second moment of group j's own distribution of the labour supply elasticity if

$$w_j^2 \quad > \quad \frac{1}{4(1-t_j)^2(1+2t_j)} sm_j \tag{15}$$

and

(ii) increasing in the respective value for any other group j' (where $j' \in \{1 \dots n\}$ and $j' \neq j$) if

$$w_{j'}^2 \quad > \quad \frac{1}{4(1-t_{j'})^2(1+2t_{j'})} sm_{j'} \quad j' \in \{1, \dots, n\}, j' \neq j, \tag{16}$$

where $sm_j \equiv \int \alpha_{ij}^2 f_j(\alpha_{ij}) \mathrm{d}\alpha_{ij}$ is the second moment of the preference distribution of group j; respectively.

Proof. Combining the first order conditions with respect to t_j and $t_{j'}$ we get:

$$\underbrace{\frac{2t_j sm_j}{(1+t_j)sm_j - 4(1-t_j)^3 w_j^2} - \frac{2t_{j'} sm_{j'}}{(1+t_{j'})sm_{j'} - 4(1-t_{j'})^3 w_{j'}^2}}_{\equiv \gamma(sm_j, sm_{j'}, t_j, t_{j'})} = 0. \tag{17}$$

The effects of a change in the second moments on the first order condition are

$$\frac{\partial \gamma}{\partial sm_j} = -\frac{8(1-t_j)t_j w_j^2}{((1+t_j)sm_j - 4(1-t_j)^3 w_j^2)^2} \leq 0 \tag{18}$$

and

$$\frac{\partial \gamma}{\partial sm'_j} = \frac{8(1-t_{j'})t_{j'} w_{j'}^2}{((1+t_{j'})sm_{j'} - 4(1-t_{j'})^3 w_{j'}^2)^2} \geq 0, \tag{19}$$

respectively. The effect of a change in the tax rates are ambiguous. More precisely,

$$\frac{\partial \gamma}{\partial t_j} = \frac{2sm_j(sm_j - 4(1-t_j)^2(1+2t_j)w_j^2)}{((1+t_j)sm_j - 4(1-t_j)^3 w_j^2)^2} \tag{20}$$

which is negative iff

$$w_j^2 \geq \frac{1}{4(1-t_j)^2(1+2t_j)} sm_j. \tag{21}$$

Analogously,

$$\frac{\partial \gamma}{\partial t_j} = \frac{2sm_j(sm_j - 4(1-t_j)^2(1+2t_j)w_j^2)}{((1+t_j)sm_j - 4(1-t_j)^3 w_j^2)^2} \tag{22}$$

is positive if

$$w_{j'}^2 \geq \frac{1}{4(1-t_{j'})^2(1+2t_{j'})} sm_{j'}. \tag{23}$$

□

On inspection of conditions (15) and (16), we see that they cannot hold if the tax rates t_j and $t_{j'}$ approximate unity. We conclude that for very high tax rates, an increase in the second moment of the preference distribution implies an increase in the optimal tax rate for the respective group. For plausible values, on the other hand, the conjecture at the beginning of this subsection appears valid.

5 Discussion and conclusion

This paper analysed how linear tax systems could efficiently be differentiated to accommodate differences between heterogenous groups. In the context of political philosophy, this touches the question of whether the principle of equality before the law ("égalité", horizontal equity) should be jettisoned in favour of increased efficiency. In the application of the principles of optimal taxation, some tax privileges have often been discussed, most notably a tax relief for females because they exhibit, on average, a higher elasticity of labour supply than males do (Alesina et al. 2007).

In keeping with existing literature on the compatibility of absolute goals with the Pareto principle, we find that the introduction of such privileges can indeed be efficient. However, *the usual focus on the first moment of the distribution of properties within groups turns out to be unacceptable.* For some symmetric distributions – the uniform and the normal distribution –, we establish by way of example and formally that the second moment also needs to be taken into account.

More generally, one would expect higher moments to come into play as soon as mean as soon as mean and variance fail to describe a distribution completely. The distribution of income, for instance, is often considered approximately log-normal. We also establish that there are *no clear-cut comparative statics*, even though we find that for moderate levels of taxation, an increase in the first and second moments for a group is likely to lead to some tax relief for that group.

The policy problem thus turned out to be a rather complex one. And this was for a *given* partitioning of taxpayers into groups, which we just posited for our analysis. In practice, the delineation of relevant groups would also be a major case of concern – in setting tax policy, do we compare all females to all males, or just single males to single females? How finely grained is the partitioning to be?

In the final analysis, this complexity may provide an instrumental rationale for the principle of horizontal equity, even on Paretian grounds: Because (i) an optimal differentiated solution may be too costly to implement and (ii) the political bickering about the partitioning of society into groups (and which groups to consider) would waste resources,[5] maybe everybody could be made better off by a rule that requires all differences in treatment to be

[5] For a similar argument, see Congleton (1997).

based on differences of individual properties, and not on statistical properties of the (various) groups that individuals belong to.[6]

We also noted that the setup of our model is incompatible with a typical assumption on the cardinalisation of preferences in the literature on heterogenous preferences in taxation, namely that differences in preferences should not be taken into account. (While, obviously, differences in abilities and initial endowment should be recognised by optimal policy.) Lockwood and Weinzierl (2012), citing Fleurbaey and Maniquet (2006), introduce this as a formal "preference neutrality" requirement. In fact, all tax rate differences in our first example (table 1) are due to diverging preferences among groups.

This brings into stark relief the kind of normative assumptions that are implicit in tax privileges such as "gender-based taxation". Formally, preference neutrality closely resembles horizontal equity in that it is a normative principle limiting the kind of information that can be taken into account in the formulation of optimal policy. As with all absolute goals, it can conflict with the Pareto principle unless an instrumental justification can be found (see lemma 1). From a practitioner's point of view, however, the latter rule has the advantage being much simpler to apply. This is because it does not require us empirically to disentangle the part of labour supply response that is due to diverging preferences from the one that is due to differences in endowments. It therefore appears not entirely implausible that horizontal equity may, in the final analysis, turn out to be the rule of thumb of choice.

References

ALESINA, A., A. ICHINO, and L. KARABARBOUNIS (2007) Gender based taxation and the division of family chores. NBER Working Paper, 13638.

BRANDER, J. A., and B. J. SPENCER (1985) Ramsey optimal two-part tariffs: the case of heterogenous groups, *Public Finance (Finances Publiques)* **40**, 335–346.

BUCHANAN, J. M., and R. D. CONGLETON (2000) Politics by principle, not interest: towards nondisciminatory democracy, *Public Choice* **102**,

[6]In the context of constitutional economics, this type of argument is in fact an old hat. See Buchanan and Congleton's (2000) "principle of generality" as well as Epstein's (1995) call for "simple rules".

377–379.

CONGLETON, R. D. (1997) Political efficiency and equal protection of the law, *Kyklos* **50**, 485–505.

EPSTEIN, R. A. (1995) Simple Rules for a Complex World. Harvard UP.

FLEURBAEY, M., and F. MANIQUET (2006) Fair income tax, *Review of Economic Studies* **73**, 55–83.

HENKING, L. (1989) Revolutions and constitutions, *Louisiana Law Review* **49**, 1023–1055.

LOCKWOOD, B., and M. C. WEINZIERL (2012) De gustibus non est taxandum: theory and evidence on preference heterogeneity and redistribution. NBER Working Paper, 17784.

MUSGRAVE, R. A. (1990) Horizontal equity, once more, *National Tax Journal* **43**, 113–122.

SANDMO, A. (1993) Optimal redistribution when tastes differ, *Finanzarchiv* **50**, 149–163.

SEN, A. (1970) The impossibility of a Paretian liberal, *Journal of Political Economy* **78**, 152–157.

STIGLITZ, J. E. (1982) Utilitarianism and horizontal equity: the case for random taxation, *Journal of Public Economics* **18**, 1–33.